IGC-08/11-3; LA-UR 08-07513
Preprint typeset using LaTeX style emulateapj v. 11/26/03

IMPACT OF STAR FORMATION INHOMOGENEITIES ON MERGER RATES AND INTERPRETATION OF LIGO RESULTS

R. O'Shaughnessy, R. Kopparapu
Center for Gravitational Wave Physics, Penn State University, 104 Davey Lab, University Park, PA 16802

K. Belczynski
Los Alamos National Lab, CCS-2/ISR-1 Group
Oppenheimer Fellow and
New Mexico State University, Dept. of Astronomy, 1320 Frenger Mall, Las Cruces, NM 88003, USA
Submitted to ApJ

ABSTRACT

Within the next decade, ground based gravitational wave detectors are in principle capable of determining the compact object merger rate per unit volume of the local universe to better than 20% with more than 30 detections. These measurements will constrain our models of stellar, binary, and star cluster evolution in the nearby present-day and ancient universe. We argue that the stellar models are sensitive to heterogeneities (in age and metallicity at least) in such a way that the predicted merger rates are subject to an additional 30-50% systematic errors unless these heterogenities are taken into account. Without adding new electromagnetic constraints on massive binary evolution or relying on more information from each merger (e.g., binary masses and spins), as few as the $\simeq 5$ merger detections could exhaust the information available in a naive comparison to merger rate predictions. As a concrete example immediately relevant to analysis of initial and enhanced LIGO results, we use a nearby-universe catalog to demonstrate that no one tracer of stellar content can be consistently used to constrain merger rates without introducing a systematic error of order $O(30\%)$ at 90% confidence (depending on the type of binary involved). For example, though binary black holes typically take many Gyr to merge, binary neutron stars often merge rapidly; different tracers of stellar content are required for these two types. More generally, we argue that theoretical binary evolution can depend sufficiently sensitively on star-forming conditions – even assuming no uncertainty in binary evolution model – that the *distribution* of star forming conditions must be incorporated to reduce the systematic error in merger rate predictions below roughly 40%. We emphasize that the degree of sensitivity to star-forming conditions depends on the binary evolution model and on the amount of relevant variation in star-forming conditions. For example, if after further comparison with electromagnetic and gravitational wave observations future population synthesis models suggest all BH-BH binary mergers occur promptly and therefore are associated with well-studied present-day star formation, the associated composition-related systematic uncertainty could be lower than the pessimistic value quoted above. Further, as gravitational wave detectors will make available many properties of each merger – binary component masses, spins, and even short GRB associations and host galaxies could be available – many detections can still be exploited to create high-precision constraints on binary compact object formation models.

Subject headings: Stars: Binaries: Close

1. INTRODUCTION

Ground based gravitational wave detector networks (LIGO, described in Abbott et al. (The LIGO Scientific Collaboration) (2003); and VIRGO, at the Virgo project website `www.virgo.infn.it`) are analyzing the results of a design-sensitivity search for the signals expected from the inspiral and merger of double compact binaries (here, NS-NS, BH-NS, BH-BH) (extending,for example, the search in Abbott et al. (The LIGO Scientific Collaboration) 2008). Sensitivity improvements in LIGO and other interferometers are expected over the next decade that will make multiple detections a near certainty. For example, based on the short lifetime of the very massive black hole X-ray binary IC-10 X-1, Bulik et al. (2008) predict that even the current generation of interferometer has a good chance of detecting a (high-mass) merger. With moderate improvements in detector sensitivity that will be in place by mid-2009 ("enhanced LIGO"), multiple detections are plausible. More conservatively, theoretical calculations which explore a wide range of still-plausible assumptions [O'Shaughnessy et al. (2008b) (PS-GRB) and O'Shaughnessy et al. (2008a) (PS-E)] predict that the advanced LIGO network is likely to detect several tens of mergers per year, allowing the merger rate per unit volume to be determined in principle to within 20%. In fact, advanced LIGO can determine the merger rate per unit volume significantly more precisely (20%; see,e.g. O'Shaughnessy et al. 2007, (PS-E2)) than measurements have constrained the star formation history of the Milky Way and the distant universe (often at least 30%; see,e.g. Pérez-González et al. 2006; Wilkins et al. 2008, and references therein),

Electronic address: oshaughn@gravity.psu.edu

against which the predictions of O'Shaughnessy et al. (2008a) and other theoretical models are normalized. As theoretical predictions can be no more precise than their input, even though a large number of merger detections are likely, advanced LIGO measurements cannot distinguish between different hypotheses about how merging binaries are produced if those merger rates differ by less than $O(30\%)$, on the basis of the number of mergers alone.

With the ability to measure both the number and properties of merging compact binaries, LIGO has long been expected to provide invaluable assistance in better-constraining hypotheses regarding compact binary formation (see,e.g., Sadowski et al. 2008, and references therein). Given a systematic error of ε in any merger rate prediction, only approximately $1/\varepsilon^2$ unique detections are needed to determine if reality is consistent with a model. Even optimistically assuming that three types of binaries can be distinguished (e.g., from their component masses) and their rates estimated at this level of accuracy, the fraction of *a priori* plausible models still consistent with the three merger rates provided by LIGO is comparatively large: $\simeq [(\log 1+\epsilon)/2]^3$ (assuming two orders of magnitude uncertainty in the a priori plausible merger rate for each type of binary), at best comparable with the miniscule fraction of parameter space needed to constrain a high-dimensional model weakly. For example, for the seven-dimensional binary evolution models compared with observations of double pulsars in O'Shaughnessy et al. (2008c), each parameter would be constrained to just $[(\log 1+\epsilon)/2]^{3/7} \simeq 0.3$ of its a priori range using merger rate estimates alone, comparable (albeit complementary) to the information provided by electromagnetic observations of double pulsars. On the contrary, had systematic errors been smaller, then detection of n binaries of each of 3 types should imply an accuracy $(2\ln 10)^{-3/7} n^{-3/14} \simeq 0.2(n/100)^{-3/14}$ in each parameter. Furthermore, since this systematic uncertainty is introduced through our lack of knowledge about the nearby and ancient universe, even though third generation detectors such as the Einstein Telescope will harvest vastly more mergers, they will be similarly limited when comparing their observed merger rates with theoretical models that rely upon existing surveys of star formation. To take full advantage of the many mergers that in-construction and third-generation instruments will detect, compact-object theorists will need to compare the distributions of binary parameters expected from theory (i.e., masses, spins) with observations.

In this paper we estimate the limiting systematic error introduced into any theoretical prediction of binary compact object merger rates through the star formation history of the universe. We furthermore explain that the relevant uncertainty is not merely overall normalization of the nearby and even distant star formation history. Instead, we argue that merger rates, particularly binary black hole merger rates, can also be sensitive to the correlated distribution of age and metallicity of their progenitor star-forming regions. Though high-precision surveys and spectral energy distribution (SED) reconstructions of galaxies may precisely determine the mean star formation rate and metallicity by the epoch of advanced LIGO, (see,e.g. Renzini 2006; Hopkins et al. 2003; Wilkins et al. 2008, and references therein), the more delicate analyses which estimate the *distribution* of star forming conditions, particularly those of low metallicity that are far more apt to produce massive black hole binaries, remain in their infancy (see,e.g. Panter et al. 2008). Given that advanced LIGO and future gravitational-wave detectors could observe mergers produced from binary stellar evolution out to as far as $z \simeq 2$ (e.g., for an optimally oriented $30+30 M_\odot$ binary black hole merger), an epoch of rapid star formation in massive galaxies, the relevant composition distribution needed to eliminate this systematic uncertainty is unlikely to be available in the near future. Equivalently, gravitational-wave interferometers will soon provide a uniquely accurate and potentially uniquely *biased* probe into the formation and evolution of high-mass stars in the early and low-metallicity universe.

1.1. *Outline and relation to prior work*

As discussed in more detail in §2, to account for local-universe inhomogeneities and to simplify the intrinsically mass-dependent results of gravitational wave searches, previous searches for gravitational wave inspiral and merger waveforms have "normalized" their result by the amount of blue light within the relevant time-averaged detection volume; see the discussion in Abbott et al. (The LIGO Scientific Collaboration) (2008) as well as the considerably more detailed presentation in Fairhurst & Brady (2007). By choosing to express results as a "merger rate per unit blue light," however, the authors limit the accuracy of any attempt to compare merger rate predictions with their observations: as emphasized above, such a comparison is helpful only to the level of accuracy that "mergers per unit blue light" can be uniquely defined. The systematic error so introduced is unlikely to seriously limit astrophysical comparisons once detections are available in the near future: initial and enhanced LIGO results, expected to have at best a handful of detections, will not reach this level of accuracy. But this composition-based systematic error is comparable to several formal uncertainties often quoted in relation to already-published upper limits and is therefore already relevant to anyone attempting to constrain their models with existing observational upper limits. In short, anyone planning on using or expressing results in this form should be aware of its limitations.

That being said, at present, gravitational-wave detectors survey only the nearby universe, where uncertainties in the distances to galaxies dominate over photometric errors (74% vs 31%, respectively; see Kopparapu et al. 2008). At best, the cumulative asymptotic luminosity can be determined only to O(10%). All of these uncertainties are comparable or greater than the uncertainty introduced into any astrophysical interpretation by assuming the number of mergers is proportional to blue light. The uncertainties discussed in this paper therefore *bound below* the accuracy of any comparison between merger rate predictions and observations.

These limiting uncertainties arise because most simple prediction (or "normalization") methods build in an implicit assumption of *homogeneity* of star-forming conditions. But because binary mergers are rare and exceptional events themselves and naturally arise more frequently from rare and exceptional conditions (e.g., old star formation or low metallicity), assuming homogene-

ity builds in systematic errors greater than the limiting uncertainty desired for advanced detectors. As outlined above and described in § 2, we argue that dividing the rate of mergers by the amount of blue light in the detection volume oversimplifies the (implicit) inverse problem: predicting how many mergers should occur given an amount of blue light. To demonstrate that other bands give different yet potentially equally relevant normalizations, we introduce a multi-band galaxy catalog for the local universe. In §3 we demonstrate that, after a starburst, different models of binary evolution and different types of binaries lead to different conclusions about the time-dependent ratio of mergers and light. We use this tool to estimate the systematic error introduced by normalizing to blue light or, more generally, any single-band normalization. For advanced detectors, galaxy catalogs will not be available. Nonetheless, as demonstrated in § 4 the gravitational wave detection rate should not be cavalierly normalized to the *mean* properties of the universe on large scales: exceptional circumstances (here, metallicity) can introduce systematic errors at least as large as the limiting uncertainty expected of advanced detectors. Though the exact magnitude of the effect cannot be determined without an equally exact theory of binary evolution, we estimate that even modestly reliable predictions could require fairly detailed input regarding the composition of the universe within reach. Finally, to clearly illustrate the effects summarized in this paper, in §5 we show that concrete, plausible predictions for a two-component universe cannot be well-modeled by a time-independent or homogeneous one.

2. LOCAL GALAXIES IN MULTIPLE BANDS

In the past, the number of mergers per unit blue light has been used to normalize the sensitivity of searches and interpret upper limits (Phinney 1991). Because blue light roughly traces current star formation and because many double neutron star mergers occur fairly soon after their progenitor binary's birth, this ratio was expected to be proportional to the fraction of massive stars that, after their second supernova, form bound double neutron stars that merge within a Hubble time. This assumption was applied widely in the theoretical (Phinney 1991; Kalogera et al. 2001) and experimental (Nutzman et al. (2004); Kopparapu et al. (2008) henceforth LGC) literature. The blue light density locally and at moderate redshift can be measured very accurately (30%, dominated by per-galaxy distance errors). Being larger than the detector's intrinsic systematic error target (15%), this measurement error has implicitly been treated in the gravitational-wave literature as the relevant systematic uncertainty on binary merger upper limits per unit star forming matter (LGC).

While adequate to zeroth logarithmic order, the traditional approach is accurate only to the degree that the universe satisfies two approximations: (i) that only present-day star formation dominates the present-day compact-object coalescence rate and (ii) that all galaxies are sufficiently similar that twice as much blue light correlates directly with twice as many mergers. In reality elliptical galaxies are expected to contribute a significant proportion of all present-day compact binary coalescence detections (de Freitas Pacheco et al. 2006), particularly from BH-BH binaries (PS-E). Because elliptical galaxies formed their stars long ago and under different star-forming conditions than the stars which produce most of the present-day blue light, normalizing coalescence detections to blue light misrepresents the relevant degrees of freedom and loses information.

To provide concrete scenarios to demonstrate that old star formation and differences between galaxies (e.g., between ellipticals and spirals; between galaxies of different formation history and metal content) can significantly influence present-day star formation and to assist in re-evaluating the systematic error associated with present-day short-range gravitational-wave observations, we construct a local-universe galaxy catalog with more information than just blue light. Rather than use a deep survey with limited sky coverage to investigate the properties of galaxies in the large distance limit, given the range relevant to the current generation of gravitational-wave interferometers we choose to extend the previous B-band catalog provided in LGC. Ideally, we would demonstrate the importance of both inhomogeneity and old star formation by using spectra of all relevant galaxies (e.g., all within $\simeq 160\,\mathrm{Mpc}$ for BH-BH mergers for initial LIGO) to reconstruct their star formation and composition histories and convolve each with an appropriate model for binary evolution. Though the situation may change as sky coverage of large-scale surveys improve, at present only photometric information is available for all galaxies out to the Virgo cluster. We have used the LEDA catalog to extract corrected U, V, and B apparent magnitudes, best distance estimates, and morphological classifications for $\simeq 38,000$ galaxies; we convert these magnitudes to luminosities using the zero-point conventions adopted in the Appendix. Though an extensive literature exists addressing methods with which to reconstruct star formation histories, metallicities, and extinctions from photometric and spectral observations (see,e.g., Gallagher et al. 1984; Kennicutt 1998; Hopkins et al. 2003; Calzetti et al. 2007; Pérez-González et al. 2006, and references therein), with so few bands we cannot reliably invert and reconstruct detailed properties of our galaxy set, even assuming the catalog uses a good IR correction to reconstruct the intrinsic U, B, V magnitude from highly obscured star formation. At best we would be limited to an $O(30\%)$ systematic error in the star-formation history reconstruction (see,e.g.,Table 4 in Pérez-González et al. 2006). We therefore work directly with the published corrected luminosities. By way of example, Figure 1 shows the cumulative luminosity versus distance for three of the bands provided in the catalog. At large distances, these three quantities match onto the average values per unit volume estimated from local-universe cosmological surveys, as discussed in the Appendix.

Ignoring differences in *when* and *how* galaxies form their stars introduces a systematic error which can be simply (under)estimated by comparing the fraction of stellar mass and blue light due to all morphological elliptical galaxies inside our detection volume (Figure 1). At large distances, elliptical galaxies account for 60% of all stellar mass but 20-40% of all light, depending on the band used; this well-known difference is extensively described in the historical and pedagogical literature (see,e.g., Cox 2000). This band-dependent difference immediately implies that any merger rate predic-

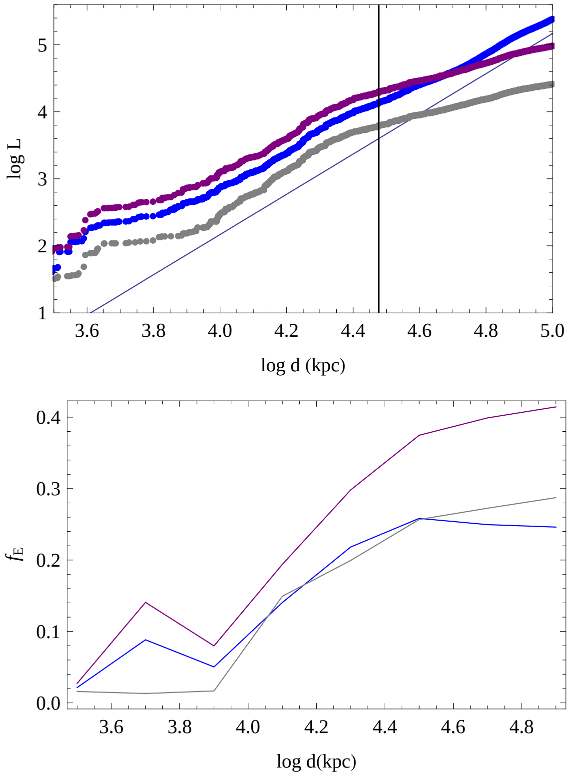

FIG. 1.— Top panel: Cumulative luminosity versus distance for three different bands (U,B,V, shown as gray, blue, and purple), shown per unit $L_{10} = 10^{10} L_{\odot}$. Also shown are (i) a thin blue line corresponding to the large-distance limit predicted from background light (see the Appendix) and (ii) a thin vertical black line at the approximate completeness limit of the survey. The two more rapid steps in cumulative luminosity correspond to the local group (at a few Mpc) and Virgo cluster (at 20 Mpc). Because our catalog does not have U,V-band light for all galaxies in our B-band complete catalog, the large-distance behavior of these bands does not approach $L \propto d^3$ at large distances. Bottom panel: Fraction of all U,B,V light inside d contributed by elliptical galaxies, versus distance; the same color scheme is used.

tion $R_{predict}$ based on multiplying the total amount of light times some merger rate per unit light must have a systematic uncertainty of order this composition uncertainty:

$$\begin{aligned} R_{predict} &= f_{light} R_{sp} + (1 - f_{light}) R_{el} \\ R_{true} &\simeq f R_{sp} + (1 - f) R_{el} \\ \delta R &= (f_{light} - f)(R_{sp} - R_{el}) \simeq \delta f O(R) \end{aligned} \tag{1}$$

where R_{sp}, R_{el} are merger rates per unit star-forming mass of spiral and elliptical galaxies respectively, f is the mass fraction in spirals, and f_{light} is some (band-dependent) light fraction in spirals. Assuming compositional or age differences cause one or the other population to dominate the present-day merger rate, the systematic uncertainty introduced by assuming the light content traces mergers should be at least $\delta f = 0.6 - 0.4 = 20\%$. As discussed in the appendix, similar uncertainties are obtained if reasonable a priori mass-to-light ratios are adopted for the two morphological types (see,e.g., Lipunov et al. 1995) or if more sophisticated estimates for M/L are adopted rather than a simple morphological classification (see,e.g., Figure 18 in Blanton & Roweis 2007). To get a better estimate and to determine what normalization is relevant – mass, light, or some combination thereof – we must model the relative proportion that past and present star formation in elliptical and spiral galaxies produce present-day mergers.

3. LIGHT AND MERGERS LAG STAR FORMATION

Since gravitational radiation drives merging binaries together exceedingly slowly, particularly for binaries with black holes which are likely not kicked close together in supernova explosions, binaries born many Gyr ago in now-old stellar populations produce a significant and occasionally overwhelming fraction of all present-day mergers (see de Freitas Pacheco et al. (2006), as well as the discussion in O'Shaughnessy et al. (2008b) and PS-E). The ratio of mergers to light in that galaxy will therefore depend not only on the star formation history of the galaxy but also on the relative rate of decay of mergers and light after a burst of star formation. The latter is significantly model- and binary type-dependent; see Figure 3 as well as the more detailed examples in PS-GRB and PS-E. No one normalization will work perfectly for all assumptions about binary evolution; for example, blue light and binary black hole mergers will rarely evolve at the same rate.

3.1. *Estimating systematic error for blue light*

The systematic error introduced by choosing to normalize to blue light instead of a quantity that decays as the desired (model-dependent) merger rate can be estimated by monte carlo over a large array of binary evolution simulations and a range of galaxy star formation histories. Specifically, if for simplicity we assume all star formation occurs in similar conditions, flux in various bands $f_{u,g,...}$ as well as the present-day total mass and star formation rate $d\rho/dt$ can all be expressed as a convolution:

$$j_X(0) = \int d\tau\ K_X(0-\tau) \frac{d\rho}{dt}(\tau) \tag{2a}$$

$$R_D(t) = \int d\tau\ K_D(t-\tau) \frac{d\rho}{dt}(\tau) \tag{2b}$$

where j_X is the luminosity density emitted per unit volume in band X and R_D is the detection rate for a network of gravitational-wave detectors of some fixed sensitivity; see O'Shaughnessy et al. (2008a) for details. The kernels K_X can be extracted from simple stellar population libraries (see Bruzual & Charlot 2003, for details and the appendix for a summary); we obtain the kernels K_D from PS-E. To incorporate the influence of old star formation on both merger rates and present-day galactic luminosities, we explored a one-parameter model motivated by studies of the star formation history of the universe:

$$\frac{d\rho}{dt}[t|\epsilon] = \dot{\rho}_o [1 + \epsilon \frac{T+t}{\tau_o} e^{-(T+t)/\tau_o}] \tag{3}$$

where ϵ is a dimensionless parameter indicating the relative importance of old star formation, $t = 0$ is the present, $t = -T$ is the big bang, $T = 13.5\,\mathrm{Gyr}$ is the age of the universe, $\tau_o = 1.5\,\mathrm{Gyr}$ is a characteristic decay time chosen so that the shape of the star formation rate (SFR) reproduces the large-redshift peak in the cosmological SFR at large ϵ (and in general resembles the overall cosmological SFR at, $\epsilon \approx 30$; see Figure 2), and

$\dot{\rho}_o = 1 M_\odot \,\mathrm{yr}^{-1}$ is a characteristic value for a galaxy's star formation rate.

Given the SFR model [Eq. 3], which depends linearly on ϵ, and the definitions of Eq. 2, all these quantities $\mathcal{X}(\epsilon)$ depend linearly on ϵ. The slopes $d\mathcal{X}/d\epsilon/\mathcal{X}(0) \equiv m_{\mathcal{X}}$ tell us the relative importance of old versus young star formation to the quantity $\mathcal{X}$; a larger m implies greater sensitivity to old star formation. To give a sense of scale, because the star formation history of the universe resembles $\rho(\epsilon \simeq 30)$, values of $m \geq 1/30 \simeq 0.03$ imply very strong dependence on old star formation: without old star formation, the quantity $\mathcal{X}$ would be at least a factor 2 smaller. To use a very concrete example, luminosities scale as

$$L_X(\epsilon) = L_X(0)(1 + m_X \epsilon) \tag{4}$$

for some m_X that we can calculate by evaluating L_X (i.e., by convolution with K_X) for any two unequal ϵ. Similarly, for each population synthesis model and each type of binary q, we can calculate m_q (e.g, q =BH-BH,BH-NS,NS-NS). The distribution of m_q (say, m_{BH-BH}, which for clarity we will denote by m_{BH}) then indicates the range of sensitivities that q binaries can have to old star formation.

Figure 3 shows our results for the distribution of m, both for the various types of light (vertical bars) and mergers (distributions, sampling a range of binary evolution assumptions). A specific set of star-forming conditions (preferred values for m), time-evolution history (preferred ϵ), and mass completely characterize that galaxy's present-day observables. To be concrete, that galaxy contributes to the cumulative blue light and number of BH-BH detections as

$$R_D : R_D(0)_{galaxy}(1 + \epsilon\, m_{BH})$$
$$L_B : L_B(0)_{galaxy}(1 + \epsilon\, m_B)$$

where the leading-order term is proportional to the mass. Because we assume all star-forming conditions are similar, the values of m are the same for all galaxies and the cumulative detection rate and light inside a volume can be found by summing over all:

$$R_D = R_D(0)(1 + \langle\epsilon\rangle\; m_{BH}) \tag{5a}$$
$$L_B = L_B(0)(1 + \langle\epsilon\rangle\; m_B) \tag{5b}$$

where $\langle\epsilon\rangle$ denotes the mass-weighted average ϵ needed to reproduce the globl star formation history of the universe and thus where for simplicity we further assume (incorrectly) that a galaxy's star formation history is independent of its mass. Therefore the ratio $\mathcal{N} = R_D/L_B$ of BH detections to blue light will explicitly depend on the model-dependent factor m_{BH}.

When blue light and mergers have exactly the same delayed response to star formation, the ratio R_D/L_B is totally independent of the star-formation history and therefore provides an excellent tool with which to constrain the underlying theory of binary evolution. In our notation, when $m_B = m_{BH}$ the two factors in Eq. 5 cancel, leading to a ratio that is independent of ϵ. More generally blue light and mergers do not mirror one another. Adopting a blue light normalization N_B by assuming $m_{BH-BH} \to m_B$ in Eq. 5 introduces a bias. To be explicit, blue light normalization assumes the once-and-for-all proportionality

$$N_B(\epsilon) \equiv \frac{R_B(\epsilon)}{L_B(\epsilon)} \to \frac{R_B(0)}{L_B(0)} = \mathcal{N}(\langle\epsilon\rangle = 0) \tag{6}$$

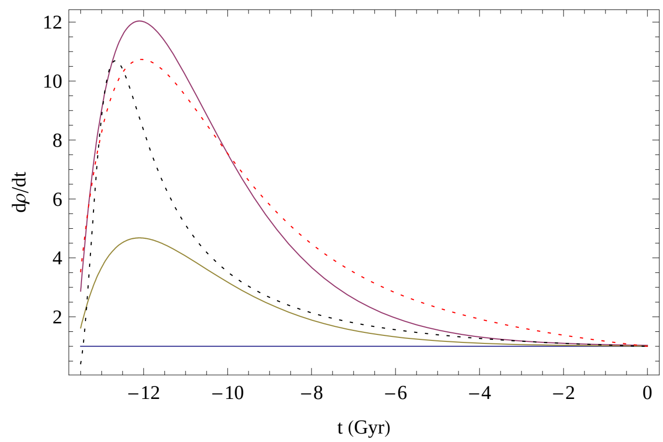

FIG. 2.— A plot of $(d\rho(t,\epsilon)/dt)/(d\rho(T)/dt)$ (Eq. 3), the one-parameter star formation history model adopted in the text, for $\epsilon = 0, 10, 30$ (blue, yellow, red, respectively). Also shown are models for $\dot{\rho}(t)/\dot{\rho}(0)$ drawn from Nagamine et al. (2006) (red, dotted) and Springel & Hernquist (2003) (black, dotted). Near $\epsilon = 30$ our one-parameter model reasonably mimics the time dependence of the star-formation history of the universe as well as of massive galaxies (Heavens et al. 2004); near the present, the model is nearly ϵ independent. The sensitivity of predictions such as Eq 2 to ϵ, as measured by "slopes" m, tell us about the relative impact of old versus young star formation.

A more detailed model that allows blue light and mergers to have different delay kernels K has a different normalization factor $\mathcal{N}(\langle\epsilon\rangle\,; m_{BH})$, which is greater than N_B by a bias factor

$$f_{bias} = \frac{\mathcal{N}(\langle\epsilon\rangle\,; m_{BH})}{N_B} = \frac{(1 + \langle\epsilon\rangle\; m_{BH})}{(1 + \langle\epsilon\rangle\; m_B)} \tag{7}$$

This bias varies depending on the model being studied. Figure 3 implies that the most-likely values for f_{bias} are between 1.4 (BH-BH) to 1.2 (NS-NS) based on a preferred value $\langle\epsilon\rangle = 30$ mentioned above and in Figure 2.

Bias isn't the most pertinent problem, however; we can always eliminate it by adopting a different convention for R_D/L_B that corresponds to the results predicted by a "typical" model. To continue with the example above, we can adopt a "typical" normalization N_{av} corresponding to Eq. (5) but with $m_{BH} \to \langle m_{BH}\rangle$. By using such a typical model, the relative bias g_{bias} between N_{av} and $\mathcal{N}$ can be much reduced:

$$\begin{aligned} g_{bias}(m_{BH-BH}) &\equiv \frac{(1 + \langle\epsilon\rangle\; m_{BH})}{(1 + \langle\epsilon\rangle\; \langle m_{BH}\rangle} \\ &= 1 + \frac{\langle\epsilon\rangle\; (m_{BH} - \langle m_{BH}\rangle)}{(1 + \langle\epsilon\rangle\; \langle m_{BH}\rangle} \end{aligned} \tag{8}$$

Nonetheless, even if we adopt the single best ratio for R_D/L_B, fluctuations between models are still sufficiently significant to significantly influence results. Specifically, the variance σ of $\ln g_{bias}$ is

$$\sigma_{\ln g_b ias} \simeq \frac{\sigma_{m_{BH}}\epsilon}{1 + \langle m_{BH}\rangle\, \epsilon} \tag{9}$$

which based on Figure 3 can be of order 1.26 (BH-BH, NS-NS) to 1.17 (NS-NS) at one standard deviation. We conclude that, depending on the type of binary involved, comparisons between theoretical models and any *single, model-independent* quantity R_D/L_B inevitably introduce

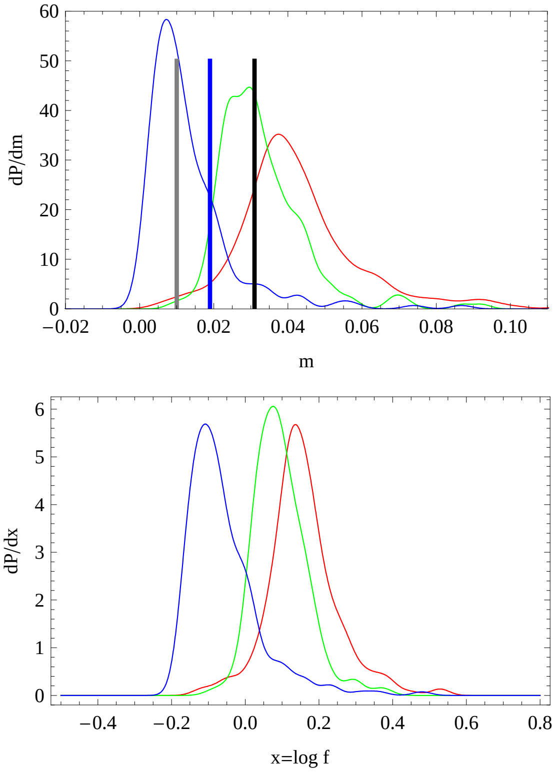

FIG. 3.— Systematic error due to optical merger tracers: Top panel shows the distribution of m for BH-BH (red), BH-NS (green) and NS-NS (blue) mergers, as well as the m values predicted for U (gray), B (blue), and V band (black) light. Note that since our simulations contain many NS-NS mergers that occur soon after formation, U-band light provides the most reliable tracer for NS-NS merger rates. Bottom panel shows the distribution of relative systematic error $x = \log \delta\mathcal{N}/\mathcal{N}$ introduced by normalizing to blue light, as predicted from the distribution of m using Eq. 7. No one band can reproduce all merger rates for all of the one-parameter star formation histories; typical systematic errors introduced by an inappropriate normalization are $O(20\%)$.

a $> 30 - 40\%$ systematic error into comparisons with binary evolution models at 90% confidence.

Is this bias really a problem?: The above calculation seems to suggest that, given a binary evolution model, blue light normalization of merger rates is biased by a known and easily-calculable factor [Eq. 7]. This correction factor can be calculated and removed post-facto, when rate predictions are compared with observations. In other words, no bias need be introduced by normalizing to any mass or light measure, so long as we can confidently relate that measure to the present-day merger rate, given assumptions about how binary mergers lag star formation of different types.

Additionally, at large distances the universe becomes homogeneous; all different light tracers become proportional, removing the need for choosing a preferred mass tracer. Normalization is most naturally made per unit volume; rate predictions are made on the basis not of galaxy models but on the star formation history of the universe (PS-GRB, PS-E and references therein). In this asymptotic case normalization is apparently unambiguous and model-dependent corrections can be reinserted later.

In fact, as we show below, independently of delay time corrections, fluctuations in composition also introduce at least as significant an uncertainty. The elliptical galaxies that host the most extreme metallicities are known to form their stars extremely early. While we could indeed correct for the contribution of old stars if all star forming conditions were similar and if the star formation history of the universe was sufficiently well-known, in the realistic heterogeneous universe systematic uncertainties in delay time and composition must both be included.

4. HETEROGENEITY AND BIAS

Star-forming conditions are known to be highly heterogeneous in time as stars gradually process metals within a galaxy, particularly for less massive galaxies which undergo extended star formation (Heavens et al. 2004). Even at the present epoch star forming conditions vary dramatically (see,e.g., Gallazzi et al. 2008; Panter et al. 2008, and references therein). Both Panter et al. (2008) and Gallazzi et al. (2008) have concluded (in their Figures 6-8 and Table 6, respectively) that nearby galaxies are likely to have metallicities Z with $\log Z/Z_\odot$ between -0.5 and 0.2. Young star-forming galaxies have an even broader range of metallicities, with $\log Z/Z_\odot$ between -1.5 and 0.2 (Figure 10 of Gallazzi et al. 2008). Though some authors have suggested even more significant differences, such as a tendency towards producing more massive stars than usual (a "top heavy IMF"; see, e.g. Hopkins & Beacom 2006, for a discussion of models and observational constraints), and though an increased number of massive stars should correspondingly increase the detection rate of compact binary coalescences, in this paper we conservatively limit attention to the more well-constrained issue of metallicity fluctuations.

The gravitational-wave detection rate R_D depends sensitively on the metallicity of the gas from which the progenitor binary stars form, as metallicity influences their structure and binary evolution. For example, observations of massive stars have demonstrated that, as expected given the larger photon cross-sections of metals over hydrogen, massive stellar winds increase significantly with more metal content (see, e.g., Vink (2008), Schröder & Cuntz (2005) and references therein). Wind loss determines the relation between initial stellar mass and final compact remnant mass of individual stars (see,e.g.,Figure 1 in Belczynski et al. 2002); as both the likelihood of a progenitor of mass M_* and the volume inside which a compact binary of chirp mass M can be observed depend sensitively on mass, metallicity fluctuations are expected to lead to significant changes in the relative likelihood and detectability of compact binary mergers. Metallicity could also influence binary evolution in other ways, such as the amount of mass lost during nonconservative mass transfer or a common-envelope phase. Unfortunately, neither observations nor theory provide an unambiguous answer for the magnitude of the effect. Theoretical methods rely on many unknown phenomenological parameters to characterize complex physical processes such as common-envelope evolution. Not only do these many unknown parameters influence merger and detection rates by orders of magnitude (Belczynski et al. 2002), they do so in a highly-correlated fashion (see,e.g.,Appendix B in O'Shaughnessy et al. 2005). Generally speaking no *single* parameter, including metallicity, produces an unam-

biguous trend everywhere in the parameter space. And equally generally the trends relevant for one type of binary (BH-BH, say) often bear little relation to the trends for other types, particularly after marginalizing over one or more other parameters.

Despite these challenges, we can fairly easily estimate the *order of magnitude* of the systematic error introduced by ignoring heterogeneity. As a first approximation we assume the composition of the universe is time-independent and estimate the present-day merger rate, averaging over the heterogeneous local universe's metallicity distribution $p(\log Z)$, as

$$\begin{aligned}\langle R_D \rangle &= \int d\log Z\, p(\log Z) \frac{dN}{dtdV_c} \\ &\times \int dM p(M|Z) V_c(M) \\ &= \int d\log Z\, p(\log Z) R_D(Z) \end{aligned} \tag{10}$$

$$V_c(M) = \frac{4\pi}{3} C_v^3 \left\langle (M/1.2M_\odot)^{15/6} \right\rangle_c \tag{11}$$

where $\log Z$ is the log of the metallicity; $dN/dtdV_c$ is the merger rate in these conditions due to *all* past star formation (and implicitly includes an integral over all time); $p(\log Z)$ is the fraction of star formation occurring in those conditions; $p(M|Z)$ is the (chirp) mass distribution of merging binaries formed due to Z; and $V_c(M)$ is the detection volume for binaries of (chirp) mass M, which we estimate using the usual power-law formula and an estimate C_V of the range at which a gravitational-wave network can detect a single double neutron star inspiral. O'Shaughnessy et al. (2008b) and O'Shaughnessy et al. (2008a) have previously performed calculations of $R_D(Z)$ for a range of metallicities and binary evolution assumptions. Based on their raw data, we estimate that the primary trend due to metallicity can be characterized by a single first-order parameter δ

$$\log R_D(Z) \simeq \log R_D(Z_\odot) + \delta \log Z/Z_\odot \tag{12}$$

defined individually for each type of binary and which allows for both the change in merger rate and in characteristic mass with metallicity. Adopting this parameter, the relative error made by ignoring heterogeneity should be of order the average value of a power law Z:

$$\langle R_D \rangle \simeq \left\langle (Z/Z_\odot)^\delta \right\rangle R_D(Z_\odot) \tag{13}$$

$$\left\langle (Z/Z_\odot)^\delta \right\rangle \simeq \frac{(Z_{max}/Z_\odot)^\delta}{\delta \ln(Z_{max}/Z_{min})} \left[1 - (Z_{min}/Z_{max})^\delta\right] \tag{14}$$

where in the second line we assume $\log Z$ is uniformly distributed between a lower and upper bound and conservatively adopt $\log Z_{max}/Z_\odot = 0.2$ and $\log Z_{min}/Z_\odot = -0.5$. Unless simulations lead to a remarkably metallicity-independent detection rate (i.e., δ is very close to zero), this expression implies that heterogeneity introduces a systematic error of order $30\% - 60\%$ for $\delta \in [-3, 6]$.[1] Though this relative change is extremely small compared to the differences between currently plausible binary evolution models for merger rates, and though this uncertainty may even be smaller than the difference between our best *StarTrack* model and reality, this error is significantly greater than the target systematic error of the LIGO analysis and greater than the eventual uncertainty of advanced LIGO measurements.

What is δ?: A worst-case estimate can be quickly extracted from the figures and results of O'Shaughnessy et al. (2008b). Merger and detection rates due to "elliptical" galaxies, in which the metallicity was varied, changed by 3 orders of magnitude (95% confidence). Assuming all this change was produced only by metallicity variation and noting metallicity varied by 0.5 in $\log Z$, we expect $|\delta| \lesssim 6$. In reality much of the observed variation is due to other parameters such as supernova kicks which strongly influence the merger rate. For example, a set of BH-NS merger rate estimates in which only Z differed suggests $\delta_{\text{BH-NS}} \simeq -2$.

Unfortunately our calculations also suggest that the derivative $d\log R_D/d\log Z$ changes depending on the binary evolution assumptions adopted. For this reason, until a model of binary evolution can be uniquely determined, the resulting heterogeneity-dependent effect is at best an unknown systematic error rather than a correctable bias. For this reason, we limit ourselves to the above estimate of *order of magnitude* of the error introduced by omitting heterogeneity in detection rate estimates. Future investigation could very well demonstrate that binary evolution is much less sensitive to metallicity than the above estimate; under these circumstances, the error introduced by ignoring heterogeneity would be much reduced.

Using strong Milky-way constraints to eliminating heterogeneity bias?: Observations of Milky Way compact binaries have long been used as stringent tests of binary evolution. For example, attempts to explain the existence of individual double white dwarfs (see, e.g., Nelemans & Tout (2005), D'Antona et al. (2006), van der Sluys et al. (2006), and references therein), binary pulsars (Wijers et al. (1992), Willems et al. (2006), and references therein), and X-ray binaries (see e.g. Podsiadlowski et al. (2002) as well as the articles and references in Lewin et al. (1995)) have constrained common-envelope evolution and the strength of supernova kicks. Similarly, the challenges of reconciling the theoretical and observed statistics of compact binary *populations* (compare, for example, Han (1998) or Belczynski et al. (2002) with Kalogera et al. (2004)) have also suggested constraints (O'Shaughnessy et al. 2008c, henceforth denoted PSC2).

Conceivably such strong constraints could uniquely determine the binary evolution model appropriate to the Milky Way. Combined with an understanding of metallicity-dependent single star evolution, we can imagine uniquely determining $R_D(Z)$. Therefore, in an ideal world, by combining $R_D(Z)$ with the metallicity distribution of the time-evolving, star-forming universe, we could produce precise merger rate predictions without ambiguity. Unfortunately, the dependence $d\log R_D/d\log Z$ of rate with metallicity *changes dramatically* between equally plausible models. Extremely strong observational constraints are required to limit attention to a small region in each parameter and therefore isolate a unique Z dependence; e.g., in PSC2 a factor x reduction in the parameter volume reduces uncertainty in each pa-

[1] Though our calculation suggests that when $\delta < 0$ the systematic error would be a factor 2, when we adopt a gaussian metallicity distribution which reproduces Table 6 of Gallazzi et al. (2008) we estimate a systematic error within the range stated.

rameter by $\simeq x^{1/7}$. Furthermore, because many of the parameters fitted through the comparison to StarTrack very plausibly could depend implicitly on metallicity, such as the strength of stellar winds, a set of parameters that reproduce the Milky Way need not reproduce other star-forming conditions. Thus strong Milky Way constraints could but need not eliminate ambiguities associated with heterogeneity.

Strong influences at low metallicity: In the above estimate we conservatively limit attention to existing populations and employ a fairly narrow metallicity distribution. Even in the local universe, very young star-forming regions can have dramatically lower metallicities and therefore contribute dramatically more mergers than allowed for above. Despite their rarity, they could dominate the merger rate. Observations of the high mass black hole in IC X-10 supports the contention that low-metallicity environments of the sort rarely considered previously could vastly dominate the present-day merger rate.

5. EXAMPLE: MULTICOMPONENT PREDICTION

In the above we have argued that an *ensemble* of binary evolution simulations may be needed to generate predictions for the distribution of star-forming conditions within the reach of future gravitational-wave detectors. A forthcoming paper by Belczynski et al will attempt to generate this ensemble in more detail, exploring the implications of many different metallicities, initial mass functions, and assumptions for binary evolution. However, to provide a concrete example that illustrates the challenges associated with heterogeneity, we construct merger rate and light predictions for a simple two-component universe following the constructive procedure in PS-GRB and PS-E. As a sufficiently realistic example involving an ensemble of metallicities is beyond the scope of this paper, we simply adopt choices for the metallicity, IMF, and binary evolution model that permit us to assemble our illustration from archival calculations of single-star spectral synthesis and massive binary evolution. Specifically, we assume our "elliptical" component has low metallicity $Z = 0.008$ and an IMF that at high masses has the fairly flat power law $d\ln N/d\ln M = p = -2.125$; our "spiral" component will have solar metallicity $Z = 0.02$ and a much steeper high-mass power law $p = -2.7$ (see Kroupa & Weidner 2003, for an explanation of this choice). The luminosity density with time is calculated according to §3 using the archived "simple stellar population" (SSP) models of Bruzual & Charlot (2003). The merger rate density with time is calculated following PS-GRB, adopting random but identical assumptions about binary evolution parameters to adopt in the `StarTrack` model. [Though these assumptions are implausible – this model assumes much higher supernova kicks ($\sigma \simeq 950 - 1000$ km/s) than are currently considered plausible – these models not only conveniently involve a metallicity that appears in the Bruzual & Charlot (2003) archives but also possess pedagogically helpful merger rate histories, as seen below.] Finally, following PS-GRB we adopt the two-component star formation history of Nagamine et al. (2006).

Figure 4 summarizes the results of this concrete example. First, as emphasized in §3, the luminosity and merger rate versus time are not simply proportional overall, both because two distinct components (ellipticals and spirals) form stars and because light and mergers each lag star formation uniquely. Second, depending on the type of merging binary of interest, different star forming conditions can dominate the merger rate. In the figure shown, spiral galaxies always dominate the BH-NS merger rate; elliptical galaxies dominate the NS-NS merger rate; and merging BH-BH binaries are produced predominantly in ellipticals early and spirals late. The unique response of stars formed in each of the two environments, combined the different time-dependent star formation histories in each environment, can produce many outcomes. Third and not indicated on the figure, the characteristic masses of merging binaries generally differs in the two components. In the case shown, the average detection-weighted chirp mass $\left\langle \mathcal{M}_c^{15/6} \right\rangle^{5/16}$ of merging BH-BH binaries in ellipticals is similar to that of spirals ($4.2 M_\odot$ in spirals, versus $5 M_\odot$ in ellipticals). On the other hand, elliptical galaxies contain noticeably less massive merging BH-NS binaries than their spiral counterparts ($1.9 M_\odot$ versus $2.6 M_\odot$).

6. CONCLUSIONS

In anticipation of an era of frequent binary coalescence detection and with the goal of divining the limiting astrophysical *measurement* uncertainties for future observations, in this paper we have examined the relevant systematic errors intrinsic to proposed *absolute* normalizations against which gravitational wave detections and upper limits can be compared. In other words, we have examined the challenges associated with comparing just the *number* of binary merger detections with predictions. We find that after a surprisingly small number of detections, either much more sophisticated models or richer data products (e.g., the observed mass distribution) will be needed to further constrain binary evolution.

For the nearby universe, relevant to initial and enhanced LIGO, we argue that the systematic error associated with using catalog-based normalizations has been understated. Though the nominal accuracy of tracers of star formation inside a volume, such as blue light as adopted in LGC, can be comparable to the systematic error target in LIGO (15%), the relevant systematic error by adopting a normalization that does not trace old mass and remains the same for all binary types and evolution models – the error introduced into any comparison between the number of detections and predictions – will be considerably larger ($\simeq 40\%$). This systematic error can be ameliorated but not eliminated by employing model-dependent normalizations. To provide a framework with which to calculate this two-band normalization, we introduced a multi-band galaxy catalog that extends the blue-light catalog presented in LGC; see Figures 1 and A6. We recommend that this catalog and approach be applied to re-evaluate the astrophysical systematic errors relevant to initial and enhanced LIGO upper limits.

Advanced detectors will probe the distant universe, for which a catalog is impractical. Though merger rates can be compared against the *average* properties of the universe, we have demonstrated that treating the universe as homogeneous will introduce *at least* a 40% systematic error, because regions of different metallicity will have different relative probabilities of producing massive

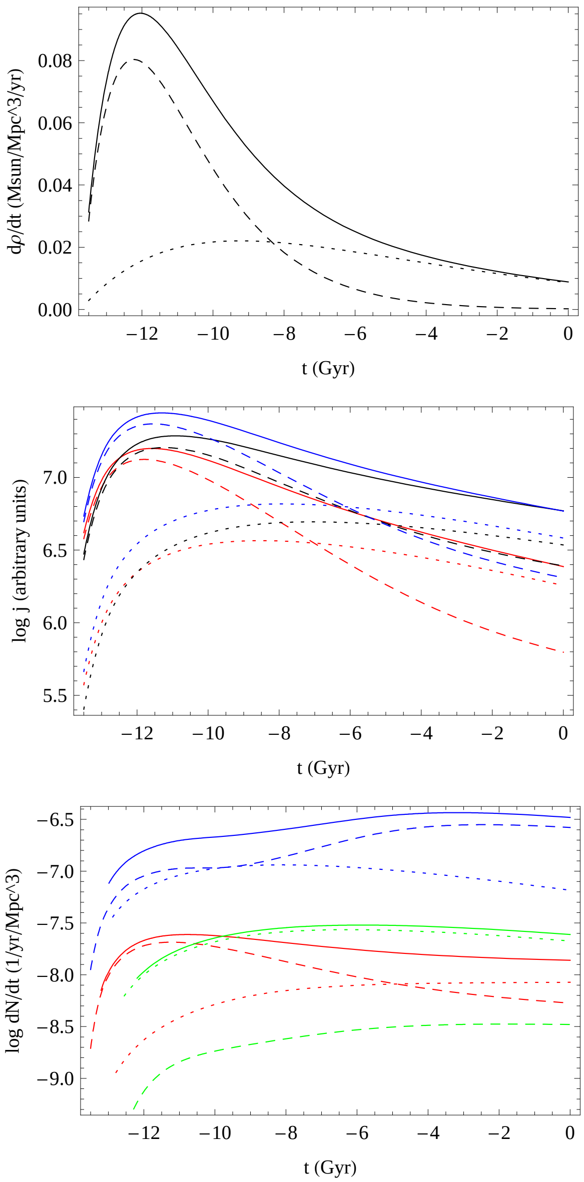

FIG. 4.— For a two-component universe with the spiral and elliptical star formation histories shown in the top panel, predictions for the time-dependent luminosity density (center panel: U (red), B (blue), V (black)) and merger rate density (bottom panel) based on Bruzual & Charlot (2003) and a pair of `StarTrack` population synthesis models for binary evolution in spiral and elliptical galaxies that adopt different IMFs and metallicity but otherwise identical parameters; see the text for details. As in Figure 2, merger rates are plotted versus time ($t = 0$ at present), where the peak at $t = -12$ Gyr corresponds to a redshift $z \sim 2$. Note the parameters adopted were chosen for convenience in illustration, not verisimilitude; for example, the very large supernova kicks assumed in this model are not consistent with observations of isolated pulsars (see,e.g. Hobbs et al. 2005). In each panel the contribution overall (solid), from spirals alone (dotted), and from ellipticals (dashed) is shown. In the bottom panel, merger rates of double neutron star (blue), double black hole (red), and black hole neutron star binaries (green) are shown. Note that merger rate densities versus time can but need not resemble light versus time and that both elliptical and spiral populations can dominate a merger rate.

merging binaries. We emphasize our estimate is conservative, assuming that the only variable in star formation is metallicity (e.g., no top-heavy IMFs or alternate modes of star formation) and that the universe was always homogeneous with a similar metallicity distribution to that observed at present. Because binary black hole detection rates in particular can be strongly influenced by metallicity variations (e.g., due to changes in the initial star-final black hole mass relation with metallicity) and because black holes are far more likely to be produced in the early universe in the epoch of peak star formation in massive galaxies undergoing rapid metallicity evolution (binary merger delays for black holes are almost always long; based on results in O'Shaughnessy et al. 2008a, the median merger delay for merging BH-BH binaries given *steady-state* star formation is $\tau_{BBH} \simeq 1 - 3\,\mathrm{Gyr}$, depending on assumptions, while for NS-NS binaries it is almost always much smaller, $\tau_{BNS} \lesssim 0.3\,\mathrm{Gyr}$), our estimate could significantly understate the relevant systematic uncertainty.

To summarize, we recommend the following: (I) When interpreting LIGO data as constraints on merger rates, unless composition distributions are explicitly incorporated into the predictive models, an additional systematic error of order 40% should be included to allow for fluctuations in composition and age between galaxies; for example, this revised uncertainty will be used in PS-E2 to explore how advanced LIGO detections might constrain binary merger models. (II) Future merger rate predictions should include metallicity evolution and distributions, to determine the most likely LIGO detection rates when low-metallicity environments are included. For example, the forthcoming paper by Belczynski et al. will explore evolutionary scenarios over ensemble of metallicities in more detail. (III) To better assess all relevant systematic errors limiting comparisons between models and theory, more observational and theoretical work is needed to constrain the distribution of fluctuations, particularly IMF fluctuations early in the universe or in clustered star formation. (IV) Finally, to provide another handle with which to constrain binary evolution, future model constraint papers should describe how to compare the detected mass distribution with highly model-dependent predictions. Given the immense computational requirements needed to both thoroughly and accurately explore the space of binary evolution models, let alone globular clusters, and the relatively modest benefits that Moore's Law provides to a monte carlo simulation sampling a high dimensional space, a careful balance must be struck between accurately pinning down predictions for each model and thoroughly exploring the model space. The parameter-dependent detection efficiency $\epsilon(D, m_1, m_2)$ and parameter-measurement-ambiguity functions provided in the gravitational-wave literature (see,e.g. Cutler & Flanagan 1994) will tell us how much we can learn about parameter distributions from LIGO and therefore determine where that balance will be struck.

When estimating systematic errors introduced by treating the universe as homogeneous, we have for simplicity assumed all mergers are produced only through binary evolution. Interactions in globular clusters are expected to be an equally critical channel for forming merging double black hole binaries; see for example

Sadowski et al. (2008) and references therein. Though we have not performed a thorough exploration of parameter space, as we were able to do for binary evolution with `StarTrack`, we expect this channel will be at least as sensitive to inhomogeneities as isolated binary evolution. More critically, this competing channel may produce mergers that are indistinguishable from binary evolution. The existence of such an unconstrained and often indistinguishable channel introduces yet another large systematic error into interpretation of binary compact object detection rates. Further study is critical, to determine not only the range of rates these models produce in a realistically heterogeneous universe but also methods with which to distinguish the randomly-oriented and equal-mass-biased mergers expected from this channel from the more aligned mergers expected from binary evolution.

We thank I. Mandel, R. Wade, A. Weinstein, and all the members of the LSC Compact Binary Coalescence group for many helpful discussions and comments over the long gestation of this paper. R.K. and R.O. were supported by National Science Foundation awards PHY 06-53462 and the Center for Gravitational Wave Physics. The Center for Gravitational Wave Physics is supported by the George A. and Margaret M. Downsbrough Endowment and by the National Science Foundation under cooperative agreement PHY 01-14375.

REFERENCES

Abbott et al. (The LIGO Scientific Collaboration). 2003, (gr-qc/0308043) [URL]
—. 2008, Phys. Rev. Lett., 77, 062002 [URL]
Belczynski, K. & et al. 2009
Belczynski, K., Kalogera, V., & Bulik, T. 2002, ApJ, 572, 407 [ADS]
Bell, E. F. & de Jong, R. S. 2001, ApJ, 550, 212 [ADS]
Blanton, M. R., Hogg, D. W., Bahcall, N. A., Brinkmann, J., Britton, M., Connolly, A. J., Csabai, I., Fukugita, M., Loveday, J., Meiksin, A., Munn, J. A., Nichol, R. C., Okamura, S., Quinn, T., Schneider, D. P., Shimasaku, K., Strauss, M. A., Tegmark, M., Vogeley, M. S., & Weinberg, D. H. 2003, ApJ, 592, 819 [ADS]
Blanton, M. R. & Roweis, S. 2007, AJ, 133, 734 [ADS]
Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000 [ADS]
Bulik, T., Belczynski, K., & Prestwich, A. 2008, (arXiv:0803.3516) [URL]
Calzetti, D., Kennicutt, R. C., Engelbracht, C. W., Leitherer, C., Draine, B. T., Kewley, L., Moustakas, J., Sosey, M., Dale, D. A., Gordon, K. D., Helou, G. X., Hollenbach, D. J., Armus, L., Bendo, G., Bot, C., Buckalew, B., Jarrett, T., Li, A., Meyer, M., Murphy, E. J., Prescott, M., Regan, M. W., Rieke, G. H., Roussel, H., Sheth, K., Smith, J. D. T., Thornley, M. D., & Walter, F. 2007, ApJ, 666, 870 [ADS]
Cox, A. N. 2000, Allen's Astrophysical Quantities, 4th edn.
Cutler, C. & Flanagan, E. 1994, 49, 2658 [ADS]
D'Antona, F., Ventura, P., Burderi, L., & Teodorescu, A. 2006, ApJ, 653, 1429 [ADS]
de Freitas Pacheco, J. A., Regimbau, T., Vincent, S., & Spallicci, A. 2006, International Journal of Modern Physics D, 15, 235 [ADS] [ADS]
Einstein Telescope. ???? [URL]
Fairhurst, S. & Brady, P. 2007, (arXiv:0707.2410) [URL]
Gallagher, III, J. S., Hunter, D. A., & Tutukov, A. V. 1984, ApJ, 284, 544 [ADS]
Gallazzi, A., Brinchmann, J., Charlot, S., & White, S. D. M. 2008, MNRAS, 383, 1439 [ADS]
Han, Z. 1998, MNRAS, 296, 1019 [ADS]
Heavens, A., Panter, B., Jimenez, R., & Dunlop, J. 2004, Nature, 428, 625 [ADS]
Hobbs, G., Lorimer, D. R., Lyne, A. G., & Kramer, M. 2005, MNRAS, 360, 974 [ADS]
Hopkins, A. M. & Beacom, J. F. 2006, ApJ, 651, 142 [ADS]
Hopkins, A. M., Miller, C. J., Nichol, R. C., Connolly, A. J., Bernardi, M., Gómez, P. L., Goto, T., Tremonti, C. A., Brinkmann, J., Ivezić, Ž., & Lamb, D. Q. 2003, ApJ, 599, 971 [ADS]
Kalogera, V., Kim, C., Lorimer, D. R., Burgay, M., D'Amico, N., Possenti, A., Manchester, R. N., Lyne, A. G., Joshi, B. C., McLaughlin, M. A., Kramer, M., Sarkissian, J. M., & Camilo, F. 2004, ApJ, 601, L179 [ADS]
Kalogera, V., Narayan, R., Spergel, D. N., & Taylor, J. H. 2001, ApJ, 556, 340 [URL]
Kennicutt, Robert C., J. 1998, Annual Review of Astronomy and Astrophysics, 36, 189 [URL]
Kopparapu, R. K., Hanna, C. R., Kalogera, V., O'Shaughnessy, R., Gonzalez, G., Brady, P. R., & Fairhurst, S. 2008, (arXiv:0706.1283) [URL]
Kroupa, P. & Weidner, C. 2003, ApJ, 598, 1076 [URL]
Lewin, van Paradijs, & van den Heuvel. 1995, X-ray Binaries (Cambridge Astrophysics)
Lipunov, V. M., Nazin, S. N., Panchenko, I. E., Postnov, K. A., & Prokhorov, M. E. 1995, A&A, 298, 677 [ADS]
Maraston, C. 1998, Monthly Notices of the Royal Astronomical Society, 300, 872 [URL]
Nagamine, K., Ostriker, J. P., Fukugita, M., & Cen, R. 2006, ApJ, 653, 881 [URL] [ADS]
Nelemans, G. & Tout, C. A. 2005, MNRAS, 356, 753 [ADS]
Nutzman, P., Kalogera, V., Finn, L. S., Hendrickson, C., & Belczynski, K. 2004, ApJ, 612, 364 [ADS]
O'Shaughnessy, R., Kalogera, V., & Belcynski, K. 2007, in preparation
—. 2008a, in preparation
O'Shaughnessy, R., Kalogera, V., & Belczynski, C. 2008b, ApJ, 675, 566+ [URL]
O'Shaughnessy, R., Kim, C., Fragos, T., Kalogera, V., & Belczynski, K. 2005, ApJ, 633, 1076 [ADS] [ADS]
O'Shaughnessy, R., Kim, C., Kalogera, V., & Belczynski, K. 2008c, ApJ, 672, 479 [URL]
Panter, B., Jimenez, R., Heavens, A. F., & Charlot, S. 2008, ArXiv e-prints, 804 [ADS]
Pérez-González, P. G., Kennicutt, Jr., R. C., Gordon, K. D., Misselt, K. A., Gil de Paz, A., Engelbracht, C. W., Rieke, G. H., Bendo, G. J., Bianchi, L., Boissier, S., Calzetti, D., Dale, D. A., Draine, B. T., Jarrett, T. H., Hollenbach, D., & Prescott, M. K. M. 2006, ApJ, 648, 987 [ADS]
Phinney, E. S. 1991, ApJ, 380, L17 [ADS]
Podsiadlowski, P., Rappaport, S., & Pfahl, E. D. 2002, ApJ, 565, 1107 [ADS]
Renzini, A. 2006, ARA&A[URL]
Sadowski, A., Belczynski, K., Bulik, T., Ivanova, N., Rasio, F. A., & O'Shaughnessy, R. 2008, ApJ, 676, 1162 [ADS]
Schröder, K.-P. & Cuntz, M. 2005, ApJ, 630, L73 [ADS]
Springel, V. & Hernquist, L. 2003, MNRAS, 339, 312 [ADS]
The VIRGO project website. ???? [URL]
van der Sluys, M. V., Verbunt, F., & Pols, O. R. 2006, A&A, 460, 209 [ADS]
Vink, J. S. 2008, New Astronomy Review, 52, 419 [ADS]
Wijers, R. A. M. J., van Paradijs, J., & van den Heuvel, E. P. J. 1992, A&A, 261, 145 [ADS]
Wilkins, S. M., Trentham, N., & Hopkins, A. M. 2008, MNRAS, 385, 687 [URL] [ADS]

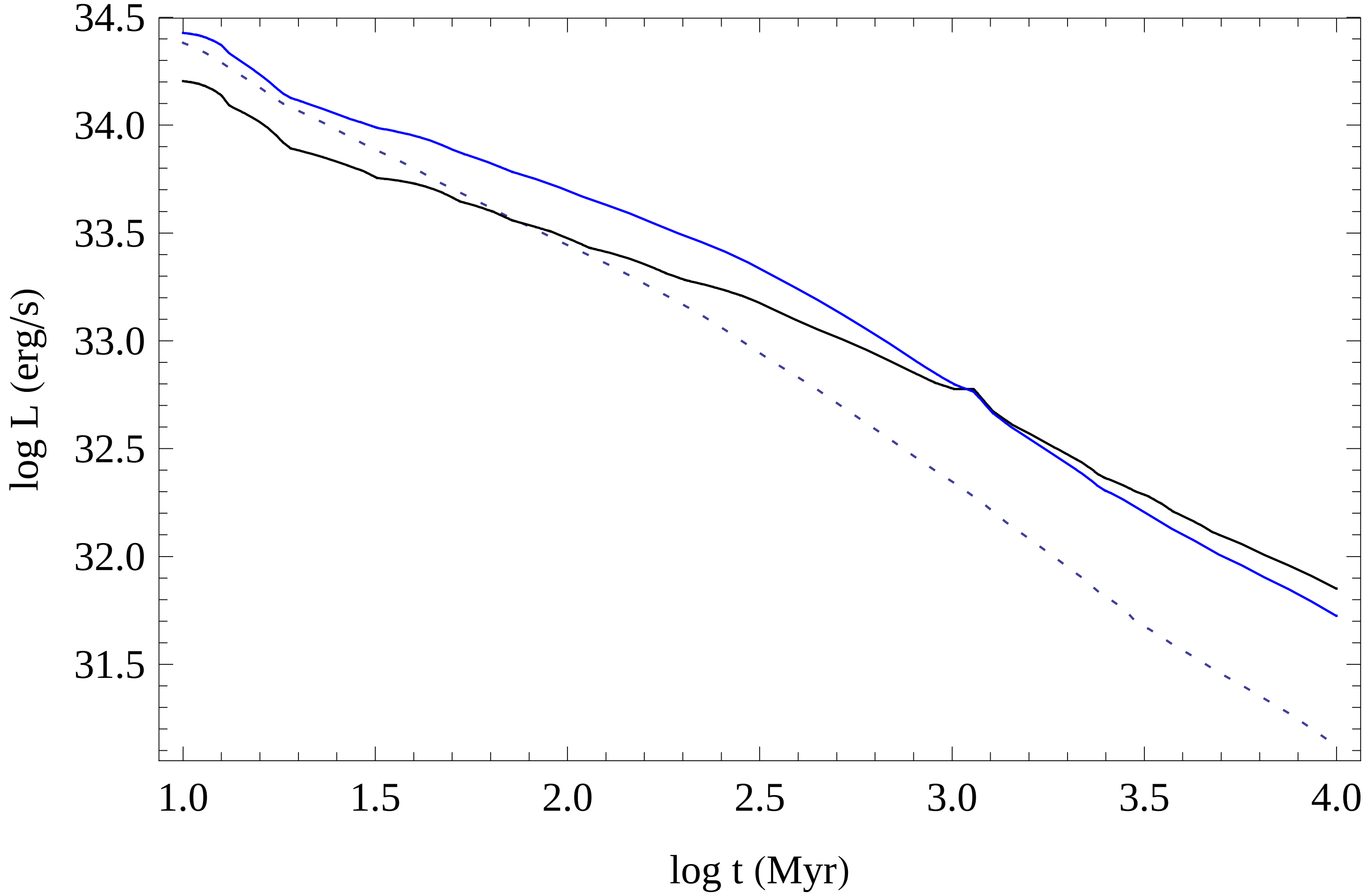

FIG. A5.— Lag between light and SFR: Total U (dotted), B (blue), and V (solid) band luminosity per $M_{\odot}$ of initial star forming mass for a starburst at time $t = 0$, drawn from the Bruzual and Charlot Bruzual & Charlot (2003) spectral synthesis libraries for solar metallicity.

Willems, B., Kaplan, J., Fragos, T., Kalogera, V., & Belczynski, K. 2006, Phys. Rev. D, 74, 043003
[ADS]

APPENDIX

PHOTOMETRY OF GALAXIES AND BACKGROUNDS

Photometric conventions: The cumulative luminosities provided in the paper are calculated from the apparent magnitudes m and distances d using the solar zero point:

$$L_X = L_{\odot,X} 10^{-0.4(m - M_{\odot,X} - 5\log(d/\,\mathrm{pc}))} \tag{A1}$$

Specifically, in this paper we adopt for blue light $L_{\odot,B} = 4.7 \times 10^{32}\,\mathrm{erg/s}$ and $M_{B,\odot} = 5.48$; for V-band light $L_{V,\odot} = 4.4 \times 10^{32}$ and $M_{V,\odot} = 4.82$; and $L_{U,\odot} = 1.7 \times 10^{32}\,\mathrm{erg/s}$ and $M_{U,\odot} = 5.66$.
Data sources for catalog: As in LGC, we use a combination of the LEDA and Tully galaxy catalogs to provide corrected distances and apparent magnitudes.
Photometric predictions from SSPs: Rather than use proportionality constants that relate the mean star formation rate to the present-day light distributions as in Kennicutt (1998), to allow for a more generic comparison we use the raw simple stellar population results provided in Bruzual & Charlot (2003) for the kernels $K_U(t, Z), K_B(t, Z), K_V(t, Z)$ that relate the star formation rate to the present-day U,B, and V luminosity densities; see Figure A5.
Mass-to-light and cumulative mass: In §2 we use the band-to-band differences in (i) the cumulative luminosity distribution and (ii) fraction from ellipticals to argue that weighting galaxies by light doubtless biases us by of order 20%. In the local universe, however, galaxies can be classified by morphology, color, or spectral information into groups with roughly similar histories and metallicities. Despite the potential advantage obtained by grouping galaxies with similar properties, unless that classification groups galaxies into sufficiently fine groups that all galaxies in that group have a similar number of present-day mergers per unit mass, the same limitations often apply; see the discussion in §3 and 4 for detailed examples.

Still, in the spirit of §3, we can introduce another observable that is linearly related to the input star formation rate: the mass (with kernel $K_M(t) = 1$). And as with different bands of light, the merger rate will be well-traced by the mass

when the merger rate kernel resembles K_M: that is, when it decays very slowly with time. As discussed in §3, fairly few binary evolution models will decay that slowly. Nonetheless, "mass normalization" (treating all star formation equally) is a meaningful and extremely complementary normalization to "blue light normalization" (emphasizing only the most recent SFR).

Depending on their stellar content, galaxies can have dramatically different stellar mass to light ratios. The literature contains several methods to estimate the relative mass content; for comparison, we adopt two methods, based on morphology and color:

- *Morphological classification*: Lipunov et al. (1995) previously used the Tully catalog and a three-component morphological classification (into elliptical, spiral, and irregular galaxies) to determine the amount of mass inside a sphere at a given radius, using the mass-to-light ratios

$$M_*/L_B(E) = 10 M_\odot/L_{\odot,B} \quad \text{(A2)}$$
$$M_*/L_B(S) = 4.5 M_\odot/L_{\odot,B} \quad \text{(A3)}$$
$$M_*/L_B(Irr) = 2 M_\odot/L_{\odot,B} \quad \text{(A4)}$$

  where any Sc or Irr galaxy is classified as irregular and young. The fraction of the cumulative "mass" distribution obtained with this estimate (Figure A6) differs to at least $O(10\%)$ from the cumulative blue luminosity.

- *Color-based M/L estimate*: Even galaxies of similar morphological type can differ substantially in their mass to light ratios (Maraston 1998; Bell & de Jong 2001). To estimate the error in the morphologically-based cumulative mass estimate described above, we use Blanton & Roweis (2007)'s Figure 18, which shows a relationship between B-V color and M_*/L_V in solar units:

$$\log_{10} M/L_V \approx 1.44(B-V)_{AB} - 0.76$$
$$-0.3(B-V)_{AB}^5 \quad \text{(A5)}$$
$$(B-V)_{AB} = (B-V)_{Vega} - 0.11 \quad \text{(A6)}$$

  where in the first line (following their figure) all magnitudes are referred to an AB magnitude system and in the second line an explicit conversion between the two magnitude systems is provided, based on their Table I.

  Based on the differences seen between the cumulative luminosity generated with this approximation and a simple morphological classification (Figure A6) or on the spread in Blanton & Roweis (2007)'s Figure 18, we expect $O(10\%)$ model-dependent uncertainty in the cumulative M and in the fraction of mass contributed from ellipticals.

Asymptotic corrected luminosity per volume: Our catalog consists of extinction-corrected ($X \equiv$)U,B, V, and far infrared (FIR) luminosities. Past the Virgo cluster, the cumulative luminosity $L_X(\leq d)$ inside a sphere of radius d should revert to a mean value

$$L_X = j_X^e \frac{4\pi}{3} d^3 \quad \text{(A7)}$$

where j_X^e is the extinction-corrected mean galactic emission per unit volume. LGC estimated the mean value $j_B^e = 1.98 \times 10^{-2} (10^{10} L_{\odot,B})\,\mathrm{Mpc}^{-3}$ by correcting the expression in Blanton et al. (2003) for the luminosity density at distances to which advanced detectors will be sensitive ($z \simeq 0.1$) by the expected amount of B-band light that should be reprocessed to FIR.

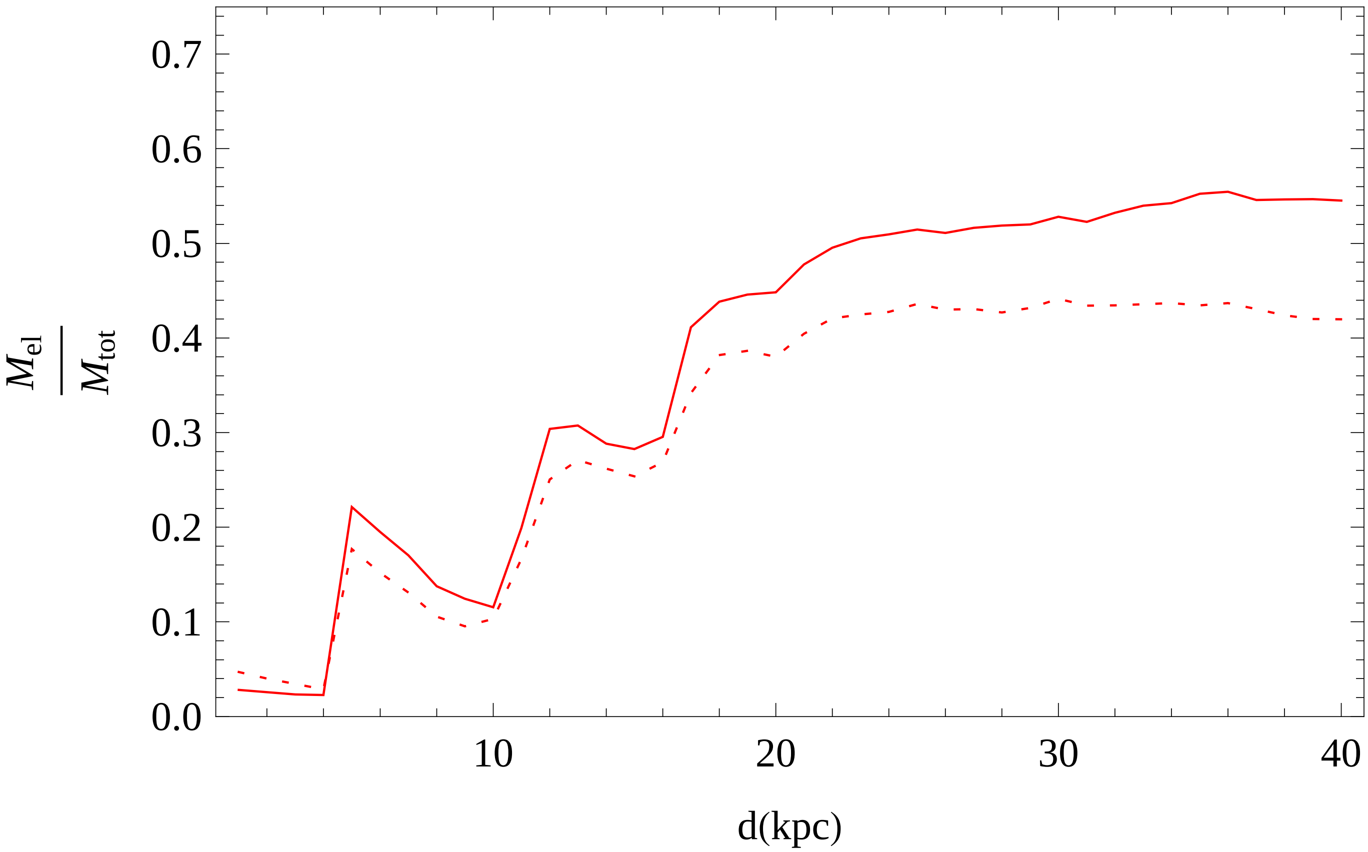

FIG. A6.— Mass fraction in ellipticals (blue) versus distance in kpc. Dotted line: Masses are estimated from B-band luminosity and the B-band mass to light ratios of Lipunov et al. (1995). Solid line: Masses are estimated using each galaxy's visible luminosity (L_V), its corrected B-V color, and and Eq. (A5), an empirical fit to the data presented in Blanton & Roweis (2007).